

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2005, 00:57:09 ; Search time 433 Seconds  
(without alignments)  
1722.604 Million cell updates/sec

Title: US-10-764-581-17  
Perfect score: 126  
Sequence: 1 catagacttaggggggga.....tgagggggaacatgatgaac 126

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	100.0	126	10	ACC43720 Nucleotid
c 2	38.4	30.5	4663	4	AB105742 Drosophil
c 3	36.4	28.9	2121	6	ABQ70526 Listeria
c 4	36	28.6	6112	4	AAL36298 Human mus
c 5	36	28.6	6112	4	AAL36297 Human mus
c 6	36	28.6	6112	8	ABX59285 cDNA enco
c 7	36	28.6	6112	8	ABX59286 cDNA enco
c 8	36	28.6	6112	12	ADJ30035 Human mus
c 9	36	28.6	6112	12	ADJ30036 Human mus
c 10	36	28.6	110000	6	ABA03041_16 Continuation (17-o
c 11	35.6	28.3	5237	6	ABQ67108 Human ang
c 12	35.4	28.1	27587	9	ADA02987 Mouse Cd2
c 13	35.4	28.1	27587	10	ADB72725 Mouse Cd2
c 14	35.4	28.1	27587	10	ADC85467 Mouse Cd2
c 15	35.4	28.1	27587	12	ADM74582 Murine ca
c 16	34.8	27.6	110000	6	ABQ69245_16 Continuation (17-o
c 17	34.8	27.6	110000	6	ABQ67195_1 Continuation (2 of
c 18	34.4	27.3	7450	6	ABK28131 DNA trans
c 19	34.4	27.3	237961	6	ABQ80552 Human Can
c 20	34.2	27.1	906	4	AAB74632 Bacillus

c 21	34.2	27.1	1619	2	AAQ71364 Expressio
c 22	34.2	27.1	1754	2	AAQ71365 Expressio
c 23	34.2	27.1	5391	6	ABK39938 Human che
c 24	34.2	27.1	5391	6	ABL32242 Human imm
c 25	34.2	27.1	8166	6	ABL33792 Human imm
c 26	33.8	26.8	5489	6	ABL34121 Human imm
c 27	33.8	26.8	15853	6	ABL70465 Chemicall
c 28	33.8	26.8	15853	6	AA61456 Human gen
c 29	33.8	26.8	23407	4	ABL06948 Drosophil
c 30	33.6	26.7	138837	13	ABD33163 Human can
c 31	33.6	26.7	202001	6	ABS52506 Human tra
c 32	33.6	26.7	202001	10	ADG46742 Human tra
c 33	33.4	26.5	3057	2	AAx59975 DNA encod
c 34	33.4	26.5	3243	2	AAx59974 DNA encod
c 35	33.4	26.5	4333	2	AAQ11563 Encodes E
c 36	33.4	26.5	15373	6	ABL32467 Human imm
c 37	33.2	26.3	2000	6	ABZ16748 Arabidops
c 38	33.2	26.3	14987	6	ABL32631 Human imm
c 39	33	26.2	747	3	AAx70266 Plasmodiu
c 40	33	26.2	8003	4	AAK81410 Human imm
c 41	33	26.2	13784	6	ABK40062 Human che
c 42	32.8	26.0	1336	3	AAc79895 Human sec
c 43	32.8	26.0	6308	6	ABL33471 Human imm
c 44	32.8	26.0	17869	6	ABK39920 Human che
c 45	32.8	26.0	17869	6	ABL32104 Human imm

ALIGNMENTS

RESULT 1  
ACC43720  
ID ACC43720 standard; DNA; 126 BP.  
XX  
AC ACC43720;  
XX  
DT 27-OCT-2003 (revised)  
DT 11-AUG-2003 (first entry)  
XX  
DE Nucleotide sequence of a fragment of the PargCo promoter.  
XX  
KW PargCo; promoter; RNA synthesis; polypeptide synthesis; cell-free system;  
KW in vitro protein synthesis; ss.  
XX  
OS Geobacillus stearothermophilus.  
XX  
PN EP1279736-A1.  
XX  
PD 29-JAN-2003.  
XX  
PF 27-JUL-2001; 2001EP-00402049.  
XX  
PR 27-JUL-2001; 2001EP-00402049.  
XX  
XX (UYNA-) UNIV NANTES.  
XX  
Sakanyan V, Snappyan M, Ghochikyan A, Lecocq F;  
WPI; 2003-373763/36.  
XX  
Synthesizing RNA or a polypeptide from a DNA template comprises adding to the reaction mixture the DNA template comprising a promoter with a UP element and encoding the desired protein and purified alpha subunit of the RNA polymerase.  
XX  
Disclosure; Fig 1; 35pp; English.

The present sequence represents a fragment of the Bacillus stearothermophilus PargCo promoter. The PargCo promoter was used to construct recombinant DNA templates to drive protein synthesis in a cell-free system in the method of the invention. The specification describes a method of RNA or polypeptide synthesis from a DNA template. The method comprises providing a cell-free system enabling RNA or polypeptide

CC synthesis from a DNA template comprising a promoter with at least one UP  
 CC element, and recovering the synthesized RNA or polypeptide. The method is  
 CC useful for synthesizing RNAs or polypeptides from a DNA template. The RNA  
 CC produced from the method is useful as an mRNA for in vitro protein  
 CC synthesis, as hybridization probes in diagnostic assays, as substrates  
 CC for analyzing processing reactions or RNA splicing, and for the  
 CC production of specific proteins of interest, such as antigens for  
 CC vaccines. (Updated on 27-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 126 BP; 49 A; 14 C; 28 G; 35 T; 0 U; 0 Other;

Query Match 100.0%; Score 126; DB 10; Length 126;  
 Best Local Similarity 100.0%; Pred. No. 1e-20;  
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CATAGACTAGGAGGGGCAAGAAAAATCCTCGAAAAATTATTAATATACATTGATT 60  
 DB 1 CATAGACTAGGAGGGGCGCAGAAAAATCCTCGAAAAATTATTAATATACATTGATT 60  
 QY 61 TTATTTTATACAGTATTATATAGAACTACATGAGGCATACGGGTGAGGGGGAACATG 120  
 DB 61 TTATTTTATACAGTATTATATAGAACTACATGAGGCATACGGGTGAGGGGGAACATG 120  
 QY 121 ATGAAC 126  
 DB 121 ATGAAC 126

RESULT 2  
 ABL05742/c  
 ID ABL05742 standard; cDNA; 4663 BP.  
 XX  
 AC ABL05742;  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 11708.  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR P-PSDB; ABB61639.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Claim 1; SEQ ID NO 11708; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 4663 BP; 1252 A; 1104 C; 1094 G; 1213 T; 0 U; 0 Other;  
 Query Match 30.5%; Score 38.4; DB 4; Length 4663;  
 Best Local Similarity 62.5%; Pred. No. 7.6;  
 Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
 QY 27 AAATCCTCGAAAAATTATTAATATATACATTTTCATTTTATTTATACAGTATTATAATGAG 86  
 DB 2341 AAATCCTCGAAAAAGTTTATATATCCATAAAACCTTTTCTTTTATAAAATGATAGAAGCTT 2282  
 QY 87 AACTACATGAGGCATACCGGTGAGGGGGAACATGAT 122  
 DB 2281 TACTACATGAGCCATCCATGATATATGTTACATGAT 2246

RESULT 3  
 ABQ70526/c  
 ID ABQ70526 standard; DNA; 2121 BP.  
 XX  
 AC ABQ70526;  
 DT 29-AUG-2003 (revised)  
 DT 29-AUG-2002 (first entry)  
 XX  
 DE Listeria monocytogenes 4b contig DNA sequence #468.  
 XX  
 KW Antibacterial; Listeria; food contamination; mutational analysis;  
 KW infection; ds.  
 XX  
 OS Listeria monocytogenes ATCC 19115.  
 XX  
 PN WO200228891-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 04-OCT-2001; 2001WO-FR003061.  
 XX  
 PR 04-OCT-2000; 2000FR-00012697.  
 XX  
 PA (INSP ) INST PASTEUR.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Kunet F, Glaser P;  
 XX  
 DR WPI; 2002-332479/37.  
 XX  
 PT New genomic sequences from Listeria species, useful for detection,  
 PT treatment and prevention of infection, also related polypeptides,  
 PT antibodies and modulators.  
 XX  
 PS Claim 14; SEQ ID NO 3339; 180pp; French.

XX The present invention relates to nucleic acid sequences (ABQ67188-  
 CC ABQ71212) from Listeria sp. The sequences are useful as probes and  
 CC primers for identification and/or detection of Listeria (e.g. as  
 CC contaminants in foods, or mutational analysis) and for analysis of gene  
 CC expression. Proteins encoded by the nucleic acid sequences can be used to  
 CC screen for compounds that modulate gene expression, replication and  
 CC pathogenicity of Listeria (potential therapeutic agents), also for  
 CC treating infections by Listeria, and are useful as immunogens in anti-  
 CC Listeria vaccines. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. (Updated  
 CC on 29-AUG-2003 to standardise OS field)  
 XX  
 SQ Sequence 2121 BP; 581 A; 449 C; 365 G; 722 T; 0 U; 4 Other;

Query Match 28.9%; Score 36.4; DB 6; Length 2121;  
 Best Local Similarity 59.8%; Pred. No. 21;  
 Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

PF	17-JAN-2001; 2001WO-US0011338.	
XX		
PR	31-JAN-2000; 2000US-0179065P.	
PR	04-FEB-2000; 2000US-0180628P.	
PR	24-FEB-2000; 2000US-0184664P.	
PR	02-MAR-2000; 2000US-0186350P.	
PR	16-MAR-2000; 2000US-0189874P.	
PR	17-MAR-2000; 2000US-0190076P.	
PR	18-APR-2000; 2000US-0198123P.	
PR	19-MAY-2000; 2000US-0205515P.	
PR	07-JUN-2000; 2000US-0209467P.	
PR	28-JUN-2000; 2000US-0214886P.	
PR	30-JUN-2000; 2000US-0215135P.	
PR	07-JUL-2000; 2000US-0216647P.	
PR	07-JUL-2000; 2000US-0216880P.	
PR	11-JUL-2000; 2000US-0217487P.	
PR	11-JUL-2000; 2000US-0217496P.	
PR	14-JUL-2000; 2000US-0218290P.	
PR	26-JUL-2000; 2000US-0220963P.	
PR	26-JUL-2000; 2000US-0220964P.	
PR	14-AUG-2000; 2000US-0224518P.	
PR	14-AUG-2000; 2000US-0224519P.	
PR	14-AUG-2000; 2000US-0225213P.	
PR	14-AUG-2000; 2000US-0225214P.	
PR	14-AUG-2000; 2000US-0225266P.	
PR	14-AUG-2000; 2000US-0225267P.	
PR	14-AUG-2000; 2000US-0225268P.	
PR	14-AUG-2000; 2000US-0225270P.	
PR	14-AUG-2000; 2000US-0225447P.	
PR	14-AUG-2000; 2000US-0225575P.	
PR	14-AUG-2000; 2000US-0225578P.	
PR	14-AUG-2000; 2000US-0225579P.	
PR	18-AUG-2000; 2000US-0226279P.	
PR	22-AUG-2000; 2000US-0226681P.	
PR	22-AUG-2000; 2000US-0226868P.	
PR	22-AUG-2000; 2000US-0227182P.	
PR	23-AUG-2000; 2000US-0227009P.	
PR	30-AUG-2000; 2000US-0228924P.	
PR	01-SEP-2000; 2000US-0229287P.	
PR	01-SEP-2000; 2000US-0229343P.	

PR	01-SEP-2000;	2000US-02293444P;
PR	01-SEP-2000;	2000US-02293455P;
PR	03-SEP-2000;	2000US-02295059P;
PR	05-SEP-2000;	2000US-02295133P;
PR	06-SEP-2000;	2000US-02295437P;
PR	06-SEP-2000;	2000US-02304388P;
PR	08-SEP-2000;	2000US-02312542P;
PR	08-SEP-2000;	2000US-02312543P;
PR	08-SEP-2000;	2000US-02312544P;
PR	08-SEP-2000;	2000US-02314133P;
PR	08-SEP-2000;	2000US-02314144P;
PR	08-SEP-2000;	2000US-02320808P;
PR	08-SEP-2000;	2000US-02320812P;
PR	12-SEP-2000;	2000US-02319688P;
PR	14-SEP-2000;	2000US-02323977P;
PR	14-SEP-2000;	2000US-02323988P;
PR	14-SEP-2000;	2000US-02323999P;
PR	14-SEP-2000;	2000US-02324000P;
PR	14-SEP-2000;	2000US-02324011P;
PR	14-SEP-2000;	2000US-02323063P;
PR	14-SEP-2000;	2000US-02323064P;
PR	14-SEP-2000;	2000US-02323065P;
PR	21-SEP-2000;	2000US-02342233P;
PR	21-SEP-2000;	2000US-02342742P;
PR	23-SEP-2000;	2000US-02343997P;
PR	25-SEP-2000;	2000US-02349988P;
PR	26-SEP-2000;	2000US-02354848P;
PR	27-SEP-2000;	2000US-02358343P;
PR	27-SEP-2000;	2000US-02358362P;
PR	29-SEP-2000;	2000US-02363277P;
PR	29-SEP-2000;	2000US-02363367P;
PR	29-SEP-2000;	2000US-02363688P;
PR	29-SEP-2000;	2000US-02363699P;
PR	29-SEP-2000;	2000US-02363700P;
PR	02-OCT-2000;	2000US-02368022P;
PR	02-OCT-2000;	2000US-02370377P;
PR	02-OCT-2000;	2000US-02370388P;
PR	02-OCT-2000;	2000US-02417855P;
PR	20-OCT-2000;	2000US-02417866P;
PR	20-OCT-2000;	2000US-02417877P;
PR	13-OCT-2000;	2000US-02399355P;
PR	13-OCT-2000;	2000US-02399372P;
PR	20-OCT-2000;	2000US-02409607P;
PR	20-OCT-2000;	2000US-02412212P;
PR	20-OCT-2000;	2000US-02417855P;
PR	08-NOV-2000;	2000US-02446177P;
PR	08-NOV-2000;	2000US-02464747P;
PR	08-NOV-2000;	2000US-02464752P;
PR	08-NOV-2000;	2000US-02464762P;
PR	08-NOV-2000;	2000US-02464777P;
PR	08-NOV-2000;	2000US-02464788P;
PR	08-NOV-2000;	2000US-02465232P;
PR	08-NOV-2000;	2000US-02465242P;
PR	08-NOV-2000;	2000US-02465252P;
PR	08-NOV-2000;	2000US-02465262P;
PR	08-NOV-2000;	2000US-02465277P;
PR	08-NOV-2000;	2000US-02465288P;
PR	08-NOV-2000;	2000US-02465322P;
PR	08-NOV-2000;	2000US-02465609P;
PR	08-NOV-2000;	2000US-02466100P;
PR	08-NOV-2000;	2000US-02466113P;
PR	08-NOV-2000;	2000US-02466133P;
PR	17-NOV-2000;	2000US-02492078P;
PR	17-NOV-2000;	2000US-02492088P;
PR	17-NOV-2000;	2000US-02492092P;
PR	17-NOV-2000;	2000US-02492100P;
PR	17-NOV-2000;	2000US-02492113P;
PR	17-NOV-2000;	2000US-02492122P;
PR	17-NOV-2000;	2000US-02492133P;



```
PR 08-SEP-2000; 2000US-0233080P.
PR 08-SEP-2000; 2000US-0233081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-023627P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the musculoskeletal system including musculoskeletal
XX cancers and also for testing and detection e.g. diagnosis.
XX
XX Example 2; SEQ ID NO 2662; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABH3087-ABH4109) associated with the musculoskeletal system useful for
XX preventing, treating or ameliorating medical conditions e.g. by protein
XX or gene therapy. The genes are isolated from a range of human tissues
XX disclosed in the specification. The nucleic acids, proteins, antibodies
XX and (ant)agonists are useful in the diagnosis, treatment and prevention
XX of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
XX adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
XX lung, or urogenital; (b) immune disorders e.g. Addison's disease,
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX arthritis and ulcerative colitis; (c) cardiovascular disorders such as
XX myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
XX cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
XX bacterial, fungal and parasitic infections. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 6112 BP; 1894 A; 905 C; 1243 G; 2070 T; 0 U; 0 Other;
XX
XX Query Match 28.6%; Score 36; DB 4; Length 6112;
XX Best Local Similarity 62.0%; Pred. No. 29;
XX Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
XX
XX QY 18 GCAAGAAAAAATCCTCGAAATATTAAATATACATTTGATTTTATTTTATACAGTAT 77
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 2667 GGAATAAAGAGTACACTACACGACTAAATGATTAATTGTATTAATGATACCTAT 2608
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 78 TATATGAGAACTACATGAGGCATACGGGTGA 109
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 2607 TATTTTAAGAATTATATATGAATGATGATAA 2576
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 6
XX ABX59285/c
XX ID ABX59285 standard; cDNA; 6112 BP.
XX
XX AC ABX59285;
XX
XX DT 26-FEB-2003 (first entry)
XX
XX cDNA encoding novel human musculoskeletal system antigen #1629.
XX
XX DE
XX
XX KW Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
```

cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;  
 KW post-operative tissue repair; limb regeneration; neuronal growth;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW AIDS-related complex; chondrocyte growth; bone regeneration;  
 KW periodontal regeneration; tissue transport; bone graft; skin aging;  
 KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;  
 KW cell growth; organ transplant; cell differentiation; body height; weight;  
 KW hair colour; eye colour; skin; percentage of adipose tissue;  
 KW pigmentation; cosmetic surgery; metabolism; biorhythm; circadian rhythm;  
 KW depression; tendency for violence; pain; reproductive capability;  
 KW hormone level; endocrine level; appetite; libido; memory; stress;  
 KW storage capability; fat content; lipid content; protein content;  
 KW carbohydrate content; vitamin content; cofactor content;  
 KW nutritional component.  
 KW  
 KW  
 OS Homo sapiens.  
 PN US2002147140-A1.  
 PP  
 XX 10-OCT-2002.  
 XX  
 XX 17-JAN-2001; 2001US-00764877.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-02114886P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244517P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR

XX (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 XX Rosen CA, Ruben SM, Barash SC;  
 XX WPI; 2003-128199/12.  
 DR  
 XX  
 XX  
 XX Isolated nucleic acid molecules encoding musculoskeletal system  
 associated polypeptides, useful for detecting disorders, e.g. cancer.  
 XX  
 XX Disclosure; SEQ ID NO 2662; 321pp; English.  
 XX  
 XX The invention describes an isolated nucleic acid molecule comprising a  
 sequence encoding musculoskeletal system associated polypeptides useful  
 for detecting disorders, e.g., cancer or cancer metastases, in animals or  
 humans. The nucleic acid: stimulates re-vascularisation of ischaemic  
 tissues associated with conditions such as thrombosis, arteriosclerosis,  
 and other cardiovascular conditions; treats wounds due to injuries,  
 burns, post-operative tissue repair, and ulcers; stimulates angiogenesis  
 and limb regeneration; stimulates neuronal growth; can treat and prevent  
 neuronal damage occurring in certain disorders or neurodegenerative  
 conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-  
 related complex; stimulates chondrocyte growth, thus they can be used to  
 enhance bone and periodontal regeneration and aid in tissue transports or  
 bone grafts; prevents skin aging due to sunburn by stimulating  
 keratinocyte growth; prevents hair loss, since RGF family members  
 activate hair-forming cells and promotes melanocyte growth; stimulates  
 growth and differentiation of hematopoietic cells and bone marrow cells  
 when used in combination with other cytokines; maintains organs before  
 transplantation or for supporting cell culture of primary tissues;  
 induces tissue of mesodermal origin to differentiate in early embryos;  
 increases or decreases the differentiation or proliferation of embryonic  
 stem cells, besides, hematopoietic lineage; modulates mammalian  
 characteristics, such as, body height, weight, hair colour, eye colour,  
 skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,  
 cosmetic surgery); modulates mammalian metabolism; changes mammal's metal  
 state or physical state by influencing biorhythms, circadian rhythms,  
 depression, tendency for violence, tolerance for pain, reproductive  
 capabilities, hormonal or endocrine levels, appetite, libido, memory, or  
 stress; increases or decreases storage capabilities, fat content, lipid,  
 protein, carbohydrate, vitamins, minerals, cofactors or other nutritional  
 components. This sequence encodes a novel human musculoskeletal system  
 antigen. Note: The sequence data for this patent did not form part of the  
 printed specification, but was obtained in electronic format directly  
 from the US patent office at  
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140  
 CC  
 XX  
 SQ Sequence 6112 BP; 1894 A; 905 C; 1243 G; 2070 T; 0 U; 0 Other;  
 Query Match 28.6%; Score 36; DB 8; Length 6112;  
 Best Local Similarity 62.0%; Pred.No. 29;  
 Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0  
 QY 18 GCAAGAAAAAATCCTCGAAAATTTATTAATATATACATTTGATTTTATTTTATACAGTAT 77  
 Db 2667 GGAAGAAAAAGAGTACACTACACGACTAAATGCGATATTTTGTATTAATGTTACCTAT 260  
 QY 78 TATAATGAGAACTACATGAGCGCATACGGGTGA 109  
 Db 2607 TATTTTGAAGAATTATATATGAATACTACTGATAA 2576  
 RESULT 7  
 ABX59286/c  
 ID ABX59286 standard; cDNA; 6112 BP.  
 XX  
 AC ABX59286;  
 XX 26-FEB-2003 (first entry)  
 DT  
 XX cDNA encoding novel human musculoskeletal system antigen #1630.



DT 20-MAY-2004 (first entry)  
XX Human musculoskeletal system-associated genomic DNA - SEQ ID 2662.  
DE musculoskeletal system; cytostatic; osteopathic; cancer; osteoporosis;  
XX gene therapy; vaccine; human; ds.  
KW Homo sapiens.  
XX US2004009488-A1.  
XX 15-JAN-2004.  
XX 13-SEP-2002; 2002US-00242515.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226273P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 12-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.



PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764877.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
XX WPI; 2004-090458/09.  
XX  
XX  
PT New nucleic acid molecule, useful for preparing a medicament for  
PT preventing, treating or ameliorating a medical condition e.g., cancer of  
PT musculoskeletal tissues or osteoporosis.  
XX  
XX Disclosure; SEQ ID NO 2662; 289pp; English.  
XX  
XX The invention relates to a novel isolated musculoskeletal system-  
CC associated nucleic acid molecule. The nucleic acid of the invention  
CC demonstrates cytostatic and osteopathic activities and may be useful for  
CC preparing a medicament for preventing, treating or ameliorating a medical  
CC condition such as cancer of the musculoskeletal tissues or osteoporosis,  
CC possibly via gene therapy or vaccine production. The current sequence is  
CC that of the human musculoskeletal system-associated genomic DNA of the  
CC invention. The current sequence is not shown within the specification per  
CC se but is available on the USPTO web-site  
CC http://seqdata.uspto.gov/sequence.html?docID=20040009488.  
XX  
XX SQ Sequence 6112 BP; 1894 A; 905 C; 1243 G; 2070 T; 0 U; 0 Other;  
  
Query Match 28.6%; Score 36; DB 12; Length 6112;  
Best Local Similarity 62.0%; Pred. No. 29;  
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
  
QY 18 GCAGAGAAAATCCTCGAAAATTAATAATATACATTGATTTATTTTATACAGTAT 77  
Db 2667 GCGAAAAAGAGTACACTAACGACTAAATGATTAATTTGTTAATGTCCTAT 2608  
  
QY 78 TATAATGAACTACATGAGGCATACGGGTGA 109  
Db 2607 TATTTAGAAATTAATAATGAACTAGTAA 2576  
  
RESULT 9  
ADJ30036/c  
ID ADJ30036 standard; DNA; 6112 BP.  
XX  
XX AC ADJ30036;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human musculoskeletal system-associated genomic DNA - SEQ ID 2663.  
XX  
XX musculoskeletal system; cytostatic; osteopathic; cancer; osteoporosis;  
KW gene therapy; vaccine; human; ds.  
XX  
XX OS Homo sapiens.  
XX  
XX US2004009488-A1.  
XX  
XX PD 15-JAN-2004.  
XX  
XX PF 13-SEP-2002; 2002US-00242515.  
XX  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0231415P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.









***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2005, 01:04:54 ; Search time 1882 Seconds  
(without alignments)  
3244.079 Million cell updates/sec

Title: US-10-764-581-17  
Perfect score: 126  
Sequence: 1 catagacttaggggggggca.....tgagggggaacatgatgaac 126

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	40.4	32.1	294250	1 AP001517	AP001517 Bacillus
C 2	39.2	31.1	123707	5 BX936455	BX936455 Zebrafish
C 3	39.2	31.1	152961	2 BX957289	BX957289 Danio rer
C 4	38.8	30.8	110000	2 PFMAL7P1_05	Continuation (6 of
C 5	38.4	30.5	2307	3 AF432229	AF432229 Drosophila
C 6	38.4	30.5	4563	6 CQ578096	CQ578096 Sequence
C 7	38.4	30.5	41909	2 AC018035	AC018035 Drosophila
C 8	38.4	30.5	51552	2 PFMAL7P1_13	Continuation (14 o
C 9	38.4	30.5	110000	2 PFMAL7P1_12	Continuation (13 o
C 10	38.4	30.5	167062	2 AC007624	AC007624 Drosophila
C 11	38.4	30.5	173702	3 AC007593	AC007593 Drosophila
C 12	38.4	30.5	333726	3 AE003789	AE003789 Drosophila
C 13	37.8	30.0	176577	9 AL157831	AL157831 Human DNA
C 14	37.6	29.8	245135	2 AC095099	AC095099 Rattus no
C 15	37.4	29.7	165038	5 BX640454	BX640454 Zebrafish
C 16	37.2	29.5	302422	1 AE017011	AE017011 Bacillus
C 17	37.2	29.5	335050	3 PFA929356	AL929356 Plasmid
C 18	37	29.4	110000	8 CR382129_21	Continuation (22 o
C 19	37	29.4	110000	8 CR382129_22	Continuation (23 o

20	37	29.4	302000	1 AP003187	AP003187 Clostridi
C 21	36.8	29.2	139820	9 AC005100	AC005100 Homo sapi
C 22	36.8	29.2	146432	2 AC079358	AC079358 Homo sapi
C 23	36.8	29.2	166860	2 AL451002	AL451002 Homo sapi
C 24	36.6	29.0	11600	1 AE013064	AE013064 Thermoana
C 25	36.6	29.0	111828	2 AC135314	AC135314 Medicago
C 26	36.6	29.0	173122	2 AC143021	AC143021 Macaca mu
C 27	36.6	29.0	348034	3 CR382400	CR382400 Plasmid
C 28	36.4	28.9	2121	6 AX416348	AX416348 Sequence
C 29	36.4	28.9	144709	2 AC120084	AC120084 Rattus no
C 30	36.4	28.9	231316	1 AE011046	AE011046 Rattus no
C 31	36.4	28.9	290507	2 AC10446	AE017327 Listeria
C 32	36.2	28.7	43666	8 AC066691	AC066691 Arabidops
C 33	36.2	28.7	109477	9 AC144983	AC144983 Xenopus t
C 34	36.2	28.7	112555	8 AC125478	AC104042 Homo sapi
C 35	36.2	28.7	128290	2 AC148995	AC125478 Medicago
C 36	36.2	28.7	160191	5 BX323069	AC148995 Medicago
C 37	36.2	28.7	183309	5 BX572641	BX323069 Zebrafish
C 38	36.2	28.7	185380	2 CR513788	BX572641 Zebrafish
C 39	36.2	28.7	186343	9 AL137248	CR513788 Danio rer
C 40	36.2	28.7	205544	2 CR847834	AL137248 Human DNA
C 41	36.2	28.7	208363	10 AL691416	CR847834 Danio rer
C 42	36.2	28.7	238543	2 CR456635	AL691416 Mouse DNA
C 43	36.2	28.7	248254	2 SPNEU1903	CR456635 Danio rer
C 44	36.2	28.7	249386	2 CR382374	AL449925 Streptoco
C 45	36.2	28.7	249386	2 CR382374	CR382374 Danio rer

## ALIGNMENTS

RESULT 1	AP001517/c	294250 bp	DNA	linear	BCT 14-JUL-2004
LOCUS	Bacillus halodurans C-125	DNA	complete genome	section 11/14.	
DEFINITION	AP001517	294250 bp	DNA	linear	BCT 14-JUL-2004
ACCESSION	AP001517	294250 bp	DNA	linear	BCT 14-JUL-2004
VERSION	AP001517.1	GI:10175500			
KEYWORDS	Bacillus halodurans C-125				
SOURCE	Bacillus halodurans C-125				
ORGANISM	Bacillus halodurans C-125				
REFERENCE	1				
AUTHORS	Takami, H.				
TITLE	Genome analysis of facultatively alkaliphilic Bacillus halodurans C-125				
JOURNAL	(in) Horikoshi, K. and Tsujii, K. (Eds.);				
REFERENCE	2				
AUTHORS	EXTREMOPHILES IN DEEP-SEA ENVIRONMENTS: 249-284;				
TITLE	Springer-Verlag (1999)				
JOURNAL	3				
REFERENCE	Takami, H. and Horikoshi, K.				
AUTHORS	Reidentification of facultatively alkaliphilic Bacillus sp. C-125				
TITLE	to Bacillus halodurans				
JOURNAL	Biosci. Biotechnol. Biochem. 63, 943-945 (1999)				
REFERENCE	4				
AUTHORS	Takami, H., Nakasone, K., Hirama, C., Takaki, Y., Masui, N., Fuji, F.,				
TITLE	Nakamura, Y. and Inoue, A.				
JOURNAL	An improved physical and genetic map of the genome of alkaliphilic				
MEDLINE	Bacillus sp. C-125				
PUBMED	Extremophiles 3 (1), 21-28 (1999)				
REFERENCE	5				
AUTHORS	Takami, H., Nakasone, K., Ogasawara, N., Hirama, C., Nakamura, Y.,				
TITLE	Masui, N., Fuji, F., Takaki, Y., Inoue, A. and Horikoshi, K.				
JOURNAL	Sequencing of three lambda clones from the genome of alkaliphilic				
MEDLINE	Bacillus sp. strain C-125				
PUBMED	Extremophiles 3 (1), 29-34 (1999)				
REFERENCE	6				
AUTHORS	Takami, H., Takaki, Y., Nakasone, K., Hirama, C., Inoue, A. and				
TITLE	Horikoshi, K.				
JOURNAL	Sequence analysis of a 32-kb region including the major ribosomal				







* be preserved.	
*	1 149984: contig of 149984 bp in length
*	149985 150084: gap of 100 bp
*	150085 152961: contig of 2877 bp in length.
FEATURES	
source	
1. .152961	
/organism="Danio rerio"	
/mol_type="genomic DNA"	
/db_xref="taxon:7955"	
/clone="CH211-232C6"	
/clone_lib="CHORI-211"	
1. .149984	
/note="assembly fragment:02029.0"	
misc_feature	
150085. .152961	
/note="assembly_fragment:02067"	
misc_feature	
150085. .152961	
/note="assembly_fragment:02067"	
ORIGIN	
Query Match 31.1%; Score 39.2; DB 2; Length 152961;	
Best Local Similarity 78.3%; Pred. No. 21;	
Matches 47; Conservative 0; Mismatches 13; Indels 0; Gaps 0;	
Qy	26 AAATCTCGAAATTAATAATACATTTGATTTTATTTTATACAGTATTATAATGA 85
Db	30149 AATATCATTTGAAGATTTTAAACAAACAATTTATGTTATTTTATAAATTTTAAATGA 30208
RESULT 4	
PFMAL7P1_05/c	
WPCOMMENT	
Sequence split into 14 fragments LOCUS PFMAL7P1 Accession AL844506	
Fragment Name Begin End	
PFMAL7P1_00 1 110000	
PFMAL7P1_01 10001 21000	
PFMAL7P1_02 20001 31000	
PFMAL7P1_03 30001 41000	
PFMAL7P1_04 40001 51000	
PFMAL7P1_05 50001 61000	
PFMAL7P1_06 60001 71000	
PFMAL7P1_07 70001 81000	
PFMAL7P1_08 80001 91000	
PFMAL7P1_09 90001 101000	
PFMAL7P1_10 100001 111000	
PFMAL7P1_11 110001 121000	
PFMAL7P1_12 120001 131000	
PFMAL7P1_13 130001 135152	
Continuation (6 of 14) of PFMAL7P1 from base 50001 (AL844506 Plasmodium falciparum 3D7	
Query Match 30.8%; Score 38.8; DB 2; Length 110000;	
Best Local Similarity 67.1%; Pred. No. 27;	
Matches 55; Conservative 0; Mismatches 27; Indels 0; Gaps 0;	
Qy	20 AAGAAAAAATCCTCGAAAAATTAATAATACATTTGATTTTATTTTATACAGTATTA 79
Db	80886 AAAAAAATAAGATAGATAATAATAATATATATATATATATATATATATATATAA 80827
Qy	80 TAATGAGAACTACATGAGGCAT 101
Db	80826 TAATTGTTACATATGAGACAT 80805
RESULT 5	
AF432229	
LOCUS	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	
AF432229	
DEFINITION	
Drosophila melanogaster flavin-containing monooxygenase FMO-2 gene, complete cds.	
ACCESSION	
AF432229	
VERSION	
AF432229.1 GI:16751749	
KEYWORDS	
Drosophila melanogaster (fruit fly)	
SOURCE	
Drosophila melanogaster	
ORGANISM	
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
LOCUS	

```
source
1. .4663
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN

Query Match          30.5%; Score 38.4; DB 6; Length 4663;
Best Local Similarity 62.5%; Pred. No. 54;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 27 AAATCTCGAAATTTAAATATACATTTGATTTTATTTTATACAGTATTATATAG 86
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2341 AAATCTGAAAGTTTATATATCCACATAAACTTTGTTTATTAATGATAGAGCTT 2282
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 AACTACATGAGCATACGGGTGAGGGGGAACATGAT 122
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2281 TACTACATGAGCCATCATGATATATGTTACATGAT 2246
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AC018035/c
LOCUS
DEFINITION Drosophila melanogaster, 41909 bp DNA linear HTG 09-DEC-1999
ACCESSION AC018035
VERSION AC018035.1 GI:6553155
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 41909)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10213107 by the submitter.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .41909
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match          30.5%; Score 38.4; DB 2; Length 41909;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 27 AAATCTCGAAATTTAAATATACATTTGATTTTATTTTATACAGTATTATATAG 86
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21273 AAATCTGAAAGTTTATATATCCACATAAACTTTGTTTATTAATGATAGAGCTT 21214
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 AACTACATGAGCATACGGGTGAGGGGGAACATGAT 122
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21213 TACTACATGAGCCATCATGATATATGTTACATGAT 21178
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
PFMAL7P1_13
WPCOMMENT
Sequence split into 14 fragments LOCUS PFMAL7P1 Accession AL844506
Fragment Name Begin End
PFMAL7P1_00 1 110000
PFMAL7P1_01 100001 210000
PFMAL7P1_02 200001 310000
PFMAL7P1_03 300001 410000
PFMAL7P1_04 400001 510000
PFMAL7P1_05 500001 610000

source
1. .41909
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match          30.5%; Score 38.4; DB 2; Length 41909;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 27 AAATCTCGAAATTTAAATATACATTTGATTTTATTTTATACAGTATTATATAG 86
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21273 AAATCTGAAAGTTTATATATCCACATAAACTTTGTTTATTAATGATAGAGCTT 21214
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 AACTACATGAGCATACGGGTGAGGGGGAACATGAT 122
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21213 TACTACATGAGCCATCATGATATATGTTACATGAT 21178
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
PFMAL7P1_13
WPCOMMENT
Sequence split into 14 fragments LOCUS PFMAL7P1 Accession AL844506
Fragment Name Begin End
PFMAL7P1_00 1 110000
PFMAL7P1_01 100001 210000
PFMAL7P1_02 200001 310000
PFMAL7P1_03 300001 410000
PFMAL7P1_04 400001 510000
PFMAL7P1_05 500001 610000
```

```
PFMAL7P1_06 600001 710000
PFMAL7P1_07 700001 810000
PFMAL7P1_08 800001 910000
PFMAL7P1_09 900001 1010000
PFMAL7P1_10 1000001 1110000
PFMAL7P1_11 1100001 1210000
PFMAL7P1_12 1200001 1310000
PFMAL7P1_13 1300001 1351552
Continuation (14 of 14) of PFMAL7P1 from base 1300001 (AL844506 Plasmodium falciparum 3)

Query Match          30.5%; Score 38.4; DB 2; Length 51552;
Best Local Similarity 64.8%; Pred. No. 38;
Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 24 AAAAATCTCGAAAATTTAAATATACATTTGATTTTATTTTATACAGTATTATAAT 83
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9554 AAAAAAAGTATATAATATAAGTACATATAGTTTATATATTATATATATCAT 9613
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 84 GAGAACTACATGAGGCATACGGGTGAGG 111
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9614 AATTATTATTGATCCACTACTTGTGATG 9641
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
PFMAL7P1_12
WPCOMMENT
Sequence split into 14 fragments LOCUS PFMAL7P1 Accession AL844506
Fragment Name Begin End
PFMAL7P1_00 1 110000
PFMAL7P1_01 100001 210000
PFMAL7P1_02 200001 310000
PFMAL7P1_03 300001 410000
PFMAL7P1_04 400001 510000
PFMAL7P1_05 500001 610000
PFMAL7P1_06 600001 710000
PFMAL7P1_07 700001 810000
PFMAL7P1_08 800001 910000
PFMAL7P1_09 900001 1010000
PFMAL7P1_10 1000001 1110000
PFMAL7P1_11 1100001 1210000
PFMAL7P1_12 1200001 1310000
PFMAL7P1_13 1300001 1351552
Continuation (13 of 14) of PFMAL7P1 from base 1200001 (AL844506 Plasmodium falciparum 3)

Query Match          30.5%; Score 38.4; DB 2; Length 110000;
Best Local Similarity 64.8%; Pred. No. 34;
Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 24 AAAAATCTCGAAAATTTAAATATACATTTGATTTTATTTTATACAGTATTATAAT 83
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109554 AAAAAAAGTATATAATATAAGTACATATAGTTTATATATTATATATATCAT 109613
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 84 GAGAACTACATGAGGCATACGGGTGAGG 111
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109614 AATTATTATTGATCCACTACTTGTGATG 109641
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AC007624/c
LOCUS
DEFINITION Drosophila melanogaster chromosome 2 clone BACR10F15 (D621) RPC1-98
10.P.15 map 42E-43A strain Y; cn bw sp, *** SEQUENCING IN PROGRESS
***, 25 unordered pieces.
ACCESSION AC007624
VERSION AC007624.5 GI:6563437
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 167062)
REFERENCE
AUTHORS Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
```



Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mada,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequiera,A., Sethi,H., Snir,E., Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.

**TITLE** Direct Submission  
**JOURNAL** Submitted (20-MAY-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA  
**COMMENT** On Mar 10, 2001 this sequence version replaced gi:7018749. Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu).

**FEATURES**  
 source  
 1..173702  
   /organism="Drosophila melanogaster"  
   /mol\_type="genomic DNA"  
   /strain="Y: cn bw sp"  
   /db\_xref="taxon:7227"  
   /chromosome="2R"  
   /map="428-43A"  
   /clone="BACR01C10 (D620)"  
   /clone\_lib="RPCI-98 (Roswell Park Cancer Institute  
   Drosophila melanogaster BAC library, partial EcoRI in  
   PBACE3.6)"

**ORIGIN**

Query Match 30.5%; Score 38.4; DB 3; Length 173702;  
 Best Local Similarity 62.5%; Pred. No. 31;  
 Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 27 AAATCCTCGAAATTTAATAATATACATTGATTTTATTTTATACAGTATTATATGAG 86  
 |||||  
 Db 102329 AAATCCTCGAAGTTTATATATCCACATAAACTTTGTTTATAAATGATAGAAGCTT 102270  
 |||||

Qy 87 AACTACATGAGCATACGGTGAGGGGACATGAT 122  
 |||||

Db 102269 TACTACATGAGCCATCATGATATATGTTACATGAT 102234  
 |||||

**RESULT 12**  
**AE003789/c**  
**LOCUS** Drosophila melanogaster chromosome 2R, section 6 of 74 of the complete sequence.  
**DEFINITION** Drosophila melanogaster chromosome 2R, section 6 of 74 of the complete sequence.  
**ACCESSION** AE003789 AE002778 AE013599  
**VERSION** AE003789.4 GI:28380699  
**KEYWORDS** Drosophila melanogaster (fruit fly)

**SOURCE** Drosophila melanogaster  
**ORGANISM** Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

**REFERENCE** 1 (bases 1 to 333736)  
**AUTHORS** Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Vandeil,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,G.L., Abail,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballou,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotter,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandrasekhar,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,I.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I.,

Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S., Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodak,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwan,C., Jalali,M., Kalush,P., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milashina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nuskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Fan,S., Pollard,J.V., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstein,G.M., Weissenbach,J., Williams,S.M., Woodagei, Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

The genome sequence of *Drosophila melanogaster*  
 Genome Biol. 287 (5461), 2185-2195 (2000)  
 20196006  
 10731132

**TITLE** 2 (bases 1 to 333736)  
**JOURNAL** Celniker,S.E., Wheeler,D.A., Krommiller,B., Carlson,J.W., Halpern,A., Patel,S., Adams,M., Champe,M., Dugan,S.P., Frise,E., Hodgson,A., George,R.A., Hoskins,R.A., Lavery,T., Muzny,D.M., Nelson,C.R., Pacleb,J.M., Park,S., Pfeiffer,B.D., Richards,S., Sodergren,E.J., Svirskaas,R., Tabor,P.E., Wan,K., Stapleton,M., Sutton,G.G., Venter,C., Weinstein,G., Scherer,S.E., Myers,E.W., Gibbs,R.A. and Rubin,G.M.  
**MEDLINE** Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence  
**PUBMED** Genome Biol. 3 (12), RESEARCH0079 (2002)  
**REFERENCE** 22426065  
**AUTHORS** 12537568

**TITLE** 3 (bases 1 to 333736)  
**JOURNAL** Misra,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S., Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E., Smith,C.D., Tupy,J.L., Whitfield,E.J., Bayraktaroglu,L., Berman,B.P., Bettencourt,B.R., Celniker,S.E., de Grey,A.D., Drysdale,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J., Shu,S.Q., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M., Rubin,G.M. and Lewis,S.E.  
**MEDLINE** Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review  
**PUBMED** Genome Biol. 3 (12), RESEARCH0083 (2002)  
**REFERENCE** 22426069  
**AUTHORS** 12537572

**TITLE** 4 (bases 1 to 333736)  
**JOURNAL** Kaminker,J.S., Bergman,C.M., Krommiller,B., Carlson,J., Svirskaas,R., Patel,S., Frise,E., Wheeler,D.A., Lewis,S.E., Rubin,G.M., Ashburner,M. and Celniker,S.E.  
**MEDLINE** The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective  
**PUBMED** Genome Biol. 3 (12), RESEARCH0084 (2002)  
**REFERENCE** 22426070  
**AUTHORS** 12537573

**TITLE** 5 (bases 1 to 333736)  
**JOURNAL** Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.  
**MEDLINE** Direct Submission  
**PUBMED** Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
**REFERENCE** 22426071  
**AUTHORS** 6 (bases 1 to 333736)

**TITLE** FlyBase  
**JOURNAL** Direct Submission  
**CONSTRM** Submitted (06-SEP-2002) University of California Berkeley, 539 Life



one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 10, constructed by the Sanger Centre Chromosome 10 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr10>

RP11-108B14 is from the library RPC1-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-108B14 The true left end of clone RP11-573G6 is at 131459 in this sequence. The true right end of clone RP11-358M4 is at 4845 in this sequence.

## FEATURES

```

source
    Location/Qualifiers
        1..176577
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="10"
            /clone="RP11-108B14"
            /clone_lib="RPC1-11.1"
            /size="159170"
            /note="CpG island"
            /evidence=not_experimental

misc_feature
    158682..159170
        /note="CpG island"
        /evidence=not_experimental

```

## ORIGIN

```

Query Match      30.0%; Score 37.8; DB 9; Length 176577;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      10  ACGGAGGCGCAGAAAAAATCTCGAAAATTTAAATATACATTGATTTTATTTA 69
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      66762 AGAGTGGCGCTAAAGTAAAGTGACAAAATACCTAAATATATATGTCTATGTTATTTA 66703

QY      70  TACAGTATTATATGAGAACTACATGAGGCATACGGTGAGGGGG 114
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      66702 TAAGTTAGACATTTAAACTACTAAAGCTACTGCTGAGAGAG 66658

```

```

RESULT 14
AC095099 LOCUS      245135 bp      DNA      linear      HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-7118, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
AC095099
VERSION     AC095099.6 GI:30467746
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Rattus norvegicus
            Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 245135)

```

```

REFERENCE
AUTHORS     Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

```

Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,K., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuheva,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakoileme,O., Okwuonu,G., Olarnpunaagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,P., Poindexter,A., Popovic,D., Primus,E., Pu,L., L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villanana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

## Direct Submission

Unpublished

2 (bases 1 to 245135)

Worley,K.C.

## Direct Submission

Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245135)

Rat Genome Sequencing Consortium.

## Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:24941115.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GCIX

Center clone name: CH230-7118

Summary Statistics

Assembly program: Atlas;

Consensus quality: 229422 bases at least Q40

Consensus quality: 232796 bases at least Q30



Consensus quality: 235239 bases at least Q20  
 Estimated insert size: 242725; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 consists of 2 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

\* 1 242788: contig of 242788 bp in length  
 \* 242789 242888: gap of unknown length  
 \* 242889 245135: contig of 2247 bp in length.

FEATURES  
 source  
 1. 245135  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-7118"  
 misc\_feature  
 1..1583  
 /note="wgs\_contig"  
 misc\_feature  
 241315..242788  
 /note="wgs\_contig"

ORIGIN  
 Query Match 29.8%; Score 37.6; DB 2; Length 245135;  
 Best Local Similarity 63.0%; Pred. No. 45;  
 Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 6 ACTTAGGAGGGCGAGAAAAATCTCGAAATATTAAATACATTTGATTATT 65  
 |||||  
 Db 186244 ACTTTGAATATATAATAAATATCCAAATAGAGAAAAAGTAGACATTTGAATTTATT 186303  
 |||||

QY 66 TTTATACAGTATTATAATGAGAACTACATGAG 97  
 |||||

Db 186304 TAGAAATATCTTATAGTCATAAATGTAAGAG 186335  
 |||||

RESULT 15  
 BX640454 165038 bp DNA linear VRT 20-MAY-2004  
 LOCUS Zebrafish DNA sequence from clone CH211-1A19 in linkage group 11,  
 DEFINITION complete sequence.  
 ACCESSION BX640454  
 VERSION BX640454.6 GI:47550478  
 KEYWORDS HTG.  
 SOURCE Danio rerio (zebrafish)  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 1 (bases 1 to 165038)  
 Hammond S.  
 Direct Submission  
 Submitted (20-MAY-2004) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On May 20, 2004 this sequence version replaced gi:46559191.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: zfish-help@sanger.ac.uk  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e. phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest, except on the rare  
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information  
 on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 Zebrafish PUC subclones occasionally display inconsistency over the  
 length of mononucleotide A/T runs and conserved TA repeats. Where  
 this is found the longest good quality representation will be  
 submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat  
 discovery system (Zhirong Bao and Sean Eddy, submitted), and those  
 beginning 'dtr' were identified by Rick Waterman (Stephen Johnson  
 lab, WashU). For further information see  
[http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml) CH211-1A19  
 is from a CHORI-211 BAC library  
 VECTOR: pTARBAC2.1.

FEATURES  
 source  
 1. 165038  
 Location/Qualifiers  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="CH211-1A19"  
 /clone\_lib="CHORI-211"

## ORIGIN

Query Match 29.7%; Score 37.4; DB 5; Length 165038;  
 Best Local Similarity 74.6%; Pred. No. 52;  
 Matches 47; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 21 AGAAAAAATCCTCGAAATATTATAATATATACATTTGATTTTATACAGTATTAT 80  
 |||||  
 Db 98697 AGAATCAAAATCTAGATTAATTAATGAATTTGATTACATATTTATGTTATTAT 98756  
 |||||

QY 81 AAT 83  
 ||  
 Db 98757 TAT 98759

Search completed: July 22, 2005, 07:46:27  
 Job time : 1893 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2005, 05:35:22 ; Search time 132 Seconds  
(without alignments)  
1561.901 Million cell updates/sec

Title: US-10-764-581-17  
Perfect score: 126  
Sequence: 1 catagactaggaggagga.....tgagggggaacatgatgaac 126

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq:  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq:  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:  
5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	36	28.6	91772	4	US-09-949-016-15568
C 2	34.2	27.1	906	3	US-09-586-935-4
C 3	34.2	27.1	1619	1	US-08-507-455-1
C 4	34.2	27.1	1754	1	US-08-507-455-2
C 5	34	27.0	43333	4	US-09-949-016-15381
C 6	33.6	26.7	202001	4	US-09-734-674-3
C 7	33.4	26.5	38559	4	US-09-949-016-13384
C 8	33.4	26.5	38559	4	US-09-949-016-13385
C 9	33.4	26.5	38559	4	US-09-949-016-13386
C 10	32.4	25.7	2625	3	US-09-453-702B-210
C 11	32.2	25.6	1168	4	US-09-949-016-4117
C 12	32.2	25.6	1272	4	US-09-673-395A-60
C 13	32.2	25.6	1309	6	5221624-4
C 14	32.2	25.6	1309	6	5221624-4
C 15	32.2	25.6	1536	4	US-09-023-655-824
C 16	32.2	25.6	1772	4	US-09-673-395A-542
C 17	32.2	25.6	5848	3	US-09-368-588-3
C 18	32.2	25.6	7859	1	US-07-854-596B-4
C 19	32.2	25.6	7859	2	US-08-450-908B-15
C 20	32.2	25.6	7859	3	US-07-982-759F-15
C 21	32.2	25.6	20598	3	US-09-593-995-10
C 22	32.2	25.6	20598	4	US-10-139-667-10
C 23	32.2	25.6	53451	4	US-09-949-016-15859
C 24	32.2	25.6	99797	4	US-09-949-016-15255
C 25	31.8	25.2	601	4	US-09-949-016-79207
C 26	31.8	25.2	3387	1	US-08-468-557-1
C 27	31.8	25.2	450395	4	US-09-949-016-15473

US-09-949-016-15568/c  
; Sequence 15568, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15568  
; LENGTH: 91772  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(91772)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15568

## ALIGNMENTS

### RESULT 1

US-09-949-016-15568/c  
; Sequence 15568, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15568  
; LENGTH: 91772  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(91772)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15568

Query Match 28.6%; Score 36; DB 4; Length 91772;  
Best Local Similarity 62.0%; Pred. No. 2.7; Mismatches 35; Indels 0; Gaps 0;

QY 18 GCAAGAAAATCTCGAAAATTTAAATATACATTTGATTTTATTTATACAGTAT 77  
DB 68218 GGAAGAAAAGACTACACTAACACGACTAAATGATTAATTTGTATTAATGACCTAT 68159  
QY 78 TATAATGAGAACTACATGAGGCATACGGGTGA 109  
DB 68158 TATTTAAGAAATATATAATGAATGACTGATAA 68127

### RESULT 2

US-09-586-935-4  
; Sequence 4, Application US/09586935  
; Patent No. 6191267  
; GENERAL INFORMATION:  
; APPLICANT: KONG, HUIMIN  
; APPLICANT: HIGGINS, LAUREN S.  
; APPLICANT: DALTON, MICHAEL

```

; APPLICANT: KUCERA, REBECCA B.
; APPLICANT: SCHILDKRAUT, IRA
; TITLE OF INVENTION: Cloning And Producing The N.BatNBI Nicking Endonuclease
; FILE REFERENCE: NEB-178
; CURRENT APPLICATION NUMBER: US/09/586,935
; CURRENT FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(903)
;
US-09-586-935-4

Query Match          27.1%; Score 34.2; DB 3; Length 906;
Best Local Similarity 58.3%; Pred. No. 3.4;
Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 17  GCGAAGAAAAAATCCTCGAAAAATTATTAATAATACATTTGATTTTATTTTATACAGTA 76
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567  GGCAAAACCAATGCACITTCATGTTCTTGATCCTCCATATGATGTGTTTATGTGATTA 636

Qy 77  TTTAATAGAACTACATAGAGGCATACGGGTGAGGGGAAACAT 119
Db 627  TCGAAATATGGAGTTTACAGTGTATTCAGACGAGAGGGAAACAT 669

```

```

RESULT 3
US-08-507-455-1/c
; Sequence 1, Application US/08507455
; Patent No. 5695961
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/507,455
; FILING DATE: 08-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: GB 9303988.1
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: MITCHARD, LEONARD C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 1498-72
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:

RESULT 4
US-08-507-455-2/c
; Sequence 2, Application US/08507455
; Patent No. 5695961
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/507,455
; FILING DATE: 08-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: GB 9303988.1
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: MITCHARD, LEONARD C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 1498-72
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1754 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 546..547
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 635..636
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 1035..1036
; FEATURE:
; NAME/KEY: misc_recomb
; LOCATION: 1411..1412
; US-08-507-455-2
Query Match 27.1%; Score 34.2; DB 1; Length 1754;
Best Local Similarity 60.0%; Pred. No. 3.8;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 1 CATGACTTAGGAGGGGCAAGAAAAATCTCGAAAAATTTATAATATACATTGTGATT 60
Db 1486 CATAGAAATATCGAATGGGAAAAAAACTGCATAAAGGCATTAAAAAGAGGGACGAT 1427
QY 61 TTATTTTATACAGTATTATAATGAGAACTACATG 95
Db 1426 TTTTTTTAAATAAAATCTTAATAATCATTAAGA 1392
RESULT 5
US-09-949-016-15381
; Sequence 15381, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15381
; LENGTH: 43333
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15381
Query Match 27.0%; Score 34; DB 4; Length 43333;
Best Local Similarity 63.4%; Pred. No. 8;
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 18 GCAAGAAAAAAATCTCGAAAAATTTATAATATACATTGTGATTTTATTTTATACAGTAT 77
Db 17141 GCAGACTCTTCGTCTTAAAAAATAAAAAAAGATTTTATTTTATAAATCT 17200
QY 78 TATAATGAGAACTACATGAGC 99
Db 17201 TATAAAATATAAAAAATTAGCC 17222
RESULT 6
US-09-949-016-15381
Query Match 26.5%; Score 33.4; DB 4; Length 38559;
Best Local Similarity 60.4%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 2 ATAGACTTAGGAGGGGCAAGAAAAATCTCGAAAAATTTATAATATACATTGTGATTT 61
Db 37868 ATATGTGGGGGTGTGGAATAGAAACATACGCAAGCATACATATTTTACTACTATTT 37927
US-09-734-674-3
; Sequence 3, Application US/09734674
; Patent No. 6498022
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; AND USES THEREOF
; FILE REFERENCE: CL001018
; CURRENT APPLICATION NUMBER: US/09/734,674
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 202001
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(202001)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-674-3
Query Match 26.7%; Score 33.6; DB 4; Length 202001;
Best Local Similarity 61.4%; Pred. No. 14;
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 35 GAAATATTATAATATACATTGTGATTTTATTTTATACAGTATTATAATGAGAACTACAT 94
Db 191642 GAGAAATATTCAATCTCTTCATTTTAAATATCAAAAATATTGTACAATAACAAAAT 191701
QY 95 GAGGCATACGGGTGAGGGGGAACATGAT 122
Db 191702 GGGGCATACATACATCAATGGAACATTAT 191729
RESULT 7
US-09-949-016-13384
; Sequence 13384, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13384
; LENGTH: 38559
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(38559)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13384
Query Match 26.5%; Score 33.4; DB 4; Length 38559;
Best Local Similarity 60.4%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 2 ATAGACTTAGGAGGGGCAAGAAAAATCTCGAAAAATTTATAATATACATTGTGATTT 61
Db 37868 ATATGTGGGGGTGTGGAATAGAAACATACGCAAGCATACATATTTTACTACTATTT 37927
```

```
Qy 62 TATTTTATACAGTATTATAATGAGAACTAC 92
||||| ||||| ||||| ||||| ||||| |||||
Db 37928 TATTATTATCCGTATATAATCTGAGACTCC 37958

RESULT 8
US-09-949-016-13385
; Sequence 13385, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13385
; LENGTH: 38559
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(38559)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13385

Query Match 26.5%; Score 33.4; DB 4; Length 38559;
Best Local Similarity 60.4%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 2 ATAGACTTAGGGGGGCAAGAAAAATCTCGAAATTTATAATATACATTTGATTT 61
||||| ||||| ||||| ||||| ||||| |||||
Db 37868 ATAATGTGGGGGTGGAAATAGAACATACGCAAGCATACATATTTTACTACTATTT 37927

Qy 62 TATTTTATACAGTATTATAATGAGAACTAC 92
||||| ||||| ||||| ||||| ||||| |||||
Db 37928 TATTATTATCCGTATATAATCTGAGACTCC 37958

RESULT 10
US-09-453-702B-210/c
; Sequence 210, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 210:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2625
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 210:
US-09-453-702B-210

Query Match 25.7%; Score 32.4; DB 3; Length 2625;
Best Local Similarity 56.6%; Pred. No. 13;
Matches 60; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 18 GCAAGAAAAAATCTCGAAAAATTTATAATATACATTTGATTTTATTTTATACAGTAT 77
||||| ||||| ||||| ||||| ||||| |||||
Db 2419 GTAAAAAGACAAATACTCTTAATAGCGGAATTTGATTTTAAATAGTAA 2360

Qy 78 TATAATGAGAACTACATGAGGCATACCGGTGAGGGGGAACATGATG 123
```

	Query Match	25.6%	Score 32.2;	DB 4;	Length 1272;
	Best Local Similarity	63.6%	Prod. No. 12;		
	Matches 49;	Conservative 0;	Mismatches 28;	Indels 0;	Gaps 0;
QY	39	ATTATTAAATATACATTGGATTTTATTTTATATACAGTATTATAATGAGAAGTACATCAGG	98		
Db	1168	ATTTTGTAGTATATAAATTTGCTTTTGTGTTTATATACGAATATAAGATTTCCCTCATTTAAT	1109		

RESULT 15  
US-09-023-655-824/c  
; Sequence 824, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.

Search completed: July 22, 2005, 08:41:12  
Job time : 137 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2005, 07:46:33 ; Search time 609 Seconds  
(without alignments)  
1312.605 Million cell updates/sec

Title: US-10-764-581-17  
Perfect score: 126  
Sequence: 1 catagacttagggaggga.....tgagggggaacatgatgaac 126

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7173243 seqs, 3172129809 residues  
Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications NA.\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
  - 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
  - 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
  - 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
  - 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
  - 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
  - 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
  - 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	100.0	126	21	US-10-764-581-17
2	36.4	28.9	2121	17	Sequence 17, Appl
3	36	28.6	6112	9	Sequence 3339, Ap
4	36	28.6	6112	9	Sequence 2662, Ap
5	36	28.6	6112	17	Sequence 2663, Ap
6	36	28.6	6112	17	Sequence 2662, Ap
7	35.6	28.3	2084	20	Sequence 2663, Ap

c	8	35.6	28.3	5237	19	US-10-433-793-138	Sequence 138, App
	9	35.4	28.1	27587	11	US-09-997-722-253	Sequence 253, App
	10	35.2	27.9	3673778	16	US-10-312-841-2	Sequence 2, Appli
c	11	35	27.8	585	13	US-10-027-632-221497	Sequence 221497,
c	12	35	27.8	585	17	US-10-027-632-221497	Sequence 221497,
c	13	34.8	27.6	495269	17	US-10-398-221-8	Sequence 8, Appli
c	14	34.8	27.6	3011208	17	US-10-398-221-2058	Sequence 2058, Ap
c	15	34.4	27.3	7450	15	US-10-240-453-5	Sequence 5, Appli
	16	34.2	27.3	237961	18	US-10-433-580-2	Sequence 2, Appli
	17	34.2	27.1	906	17	US-10-276-289-4	Sequence 4, Appli
	18	34.2	27.1	5391	15	US-10-311-455-215	Sequence 215, App
c	19	34.2	27.1	5391	17	US-10-257-166-19	Sequence 19, Appl
c	20	34.2	27.1	8166	15	US-10-311-455-1765	Sequence 1765, Ap
c	21	33.8	26.8	5489	15	US-10-311-455-2094	Sequence 2094, Ap
c	22	33.8	26.8	15853	17	US-10-221-613-421	Sequence 421, App
c	23	33.6	26.7	138837	19	US-10-322-281-146	Sequence 146, App
	24	33.6	26.7	202001	9	US-09-734-674-3	Sequence 3, Appli
	25	33.6	26.7	202001	14	US-10-274-990-3	Sequence 3, Appli
	26	33.6	26.7	202001	24	US-11-061-825-3	Sequence 3, Appli
	27	33.4	26.5	15373	15	US-10-311-455-440	Sequence 440, App
	28	33.2	26.3	2000	9	US-09-938-842A-4553	Sequence 4553, Ap
	29	33.2	26.3	2000	11	US-09-938-842A-4553	Sequence 604, App
c	30	33.2	26.3	14987	15	US-10-311-455-604	Sequence 5, Appli
c	31	33.2	26.3	380963	22	US-10-737-082-5	Sequence 5, Appli
	32	33.2	26.3	380963	22	US-10-785-790-5	Sequence 5, Appli
	33	33	26.2	201	20	US-10-719-993-20393	Sequence 20393, A
c	34	33	26.2	201	20	US-10-719-993-50769	Sequence 50769, A
c	35	33	26.2	201	21	US-10-741-600-40227	Sequence 40227, A
c	36	33	26.2	201	21	US-10-741-600-67698	Sequence 67698, A
	37	33	26.2	777	13	US-10-027-632-32741	Sequence 32741, A
	38	33	26.2	777	17	US-10-027-632-32741	Sequence 32741, A
	39	33	26.2	13784	17	US-10-257-166-144	Sequence 144, App
c	40	33	26.2	24438	20	US-10-719-993-7017	Sequence 7017, Ap
c	41	33	26.2	24438	21	US-10-741-600-17939	Sequence 17939, A
c	42	33	26.2	1980090	20	US-10-719-993-6815	Sequence 6815, Ap
c	43	33	26.2	1980090	21	US-10-741-600-17676	Sequence 17676, A
c	44	32.8	26.0	6308	15	US-10-311-455-1444	Sequence 1444, Ap
c	45	32.8	26.0	17869	15	US-10-311-455-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1  
US-10-764-581-17  
; Sequence 17, Application US/10764581  
; Publication No. US20050032086A1  
; GENERAL INFORMATION:  
; APPLICANT: UNIVERSITE DE NANTES  
; TITLE OF INVENTION: IMPROVED METHODS OF RNA AND PROTEIN SYNTHESIS  
; FILE REFERENCE: B4949AB  
; CURRENT APPLICATION NUMBER: US/10/764,581  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: PCT/EP 02/09423  
; PRIOR FILING DATE: 2002-07-26  
; PRIOR APPLICATION NUMBER: EP 01402049.9  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 126  
; TYPE: DNA  
; ORGANISM: Bacillus stearothermophilus  
US-10-764-581-17

Query Match 100.0%; Score 126; DB 21; Length 126;  
Best Local Similarity 100.0%; Pred. No. 3.8e-20;  
Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CATAGACTTAGGGAGGCGAAGAAAAAATCTCGAAAAATTTAAATATACATTGATT 60  
Db 1 CATAGACTTAGGGAGGCGAAGAAAAAATCTCGAAAAATTTAAATATACATTGATT 60

QY 61 TTATTTTATACAGTATTATATGAGAACTACATGAGGCATACGGGTGAGGGGGAACATG 120  
|||||  
Db 61 TTATTTTATACAGTATTATATGAGAACTACATGAGGCATACGGGTGAGGGGGAACATG 120  
|||||  
QY 121 ATGAAC 126  
|||||  
Db 121 ATGAAC 126

RESULT 2  
US-10-398-221-3339/c  
; Sequence 3339, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederik  
; APPLICANT: GLASER, Philippe  
; TITLE OF INVENTION: Listeria innocua, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; CURRENT FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3339  
; LENGTH: 2121  
; TYPE: DNA  
; ORGANISM: Listeria monocytogenes 4b  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)\_(end)  
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u  
US-10-398-221-3339

Query Match 28.9%; Score 36.4; DB 17; Length 2121;  
Best Local Similarity 59.8%; Pred. No. 70;  
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 24 AAAAAATCCTCGAAAAATTTAAATATATACATTGATTTTATTTTATATACAGTATTATAAT 83  
|||||  
Db 219 AAAATAGTTTAAATATTAAATAAAAAACCATTTTGACATTATAAATATACATCATTATAAT 160  
|||||  
QY 84 GAGAACTACATGAGGCATACGGGTGAGGGGGAACATGATGAA 125  
|||||  
Db 159 GAAAGTATCATTTTAATTAATTAACGGGAGGGTTAGGATGAA 118  
|||||

RESULT 3  
US-09-764-877-2662/c  
; Sequence 2662, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2662  
; LENGTH: 6112  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-2662

Query Match 28.6%; Score 36; DB 9; Length 6112;  
Best Local Similarity 62.0%; Pred. No. 1.1e+02;  
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 18 GCAAGAAAAAAATCCTCGAAAAATTTATAATATACATTGATTTTATTTTATACAGTAT 77

Db 2667 GGAAGAAAAAGTAGTACACTTAACACGCACTAAATGGATATTAAATTTGTATAAATGTACCCAT 2608  
|||||  
QY 78 TATAATGAGAACTACATGAGGCATACGGGTGA 109  
|||||  
Db 2607 TATTTTAAGAATTATATAATGAATGACTGATAA 2576  
|||||

RESULT 4  
US-09-764-877-2663/c  
; Sequence 2663, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2663  
; LENGTH: 6112  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-2663

Query Match 28.6%; Score 36; DB 9; Length 6112;  
Best Local Similarity 62.0%; Pred. No. 1.1e+02;  
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 18 GCAAGAAAAAAATCCTCGAAAAATTTATAATATACATTGATTTTATTTTATACAGTAT 77  
|||||  
Db 2667 GGAAGAAAAAGTAGTACACTTAACACGCACTAAATGGATATTAAATTTGTATAAATGTACCCAT 2608  
|||||  
QY 78 TATAATGAGAACTACATGAGGCATACGGGTGA 109  
|||||  
Db 2607 TATTTTAAGAATTATATAATGAATGACTGATAA 2576  
|||||

RESULT 5  
US-10-242-515-2662/c  
; Sequence 2662, Application US/10242515  
; Publication No. US20040009488A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005C1  
; CURRENT APPLICATION NUMBER: US/10/242,515  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 09/764,877  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0



```

; SEQ ID NO 2662
; LENGTH: 6112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-2662

Query Match      28.6%; Score 36; DB 17; Length 6112;
Best Local Similarity 62.0%; Pred. No. 1.1e+00;
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy      18  GCAGAGAAAAAATCCTCGAAAAATATTATTAATAATATACATTGATTTTATTTTATATACAGTAT 77
      |||
Db      2667  GGAAAAAAGAGTACATCAACGACTAATATGGATATTAATTTGTATAAATGTACCCCTAT 2608

Qy      78  TATTAATCAGAACTCATGAGGATACGGGTGA 109
      |||
Db      2607  TATTTTGAAGATTATATAATGAATCTGATAA 2576

```

## RESULT 6

```

US-10-242-515-2663/c
; Sequence 2663, Application US/10242515
; Publication No. US2004009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242.515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2663
; LENGTH: 6112
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-2663

```

	Query Match	28.6%	Score 36;	DB 17;	Length 6112;
	Best Local Similarity	62.0%;	Pred. No. 1.1e+02;		
	Matches 57;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;
Qy	18	GCAGAGAAAAATCCTCGAAATTTATTAAATATACATTTGATTTTATTTTATACAGTAT	77		
Db	2667	GGAAAAAAGAGTACACTAACACACACTAAATGGATATTAAATTTGTATAAATGTACCCAT	2608		
Qy	78	TATAATCGAGAACTACATGAGGCATACGGTGA	109		
Db	2607	TATTTTAAAGAAATTATAAATGAATPACTGATAA	2576		

## RESULT 7

US-10-425-115-166581/c

```

: Sequence 166581, Application US/10425115
: Publication No. US20040214272A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(53222)B
: CURRENT APPLICATION NUMBER: US/10/425,115
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 369326
: SEQ ID NO 166581
: LENGTH: 2084
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: MRT4577_83501C.1
: US-10-425-115-166581

```

	Query Match	28.3%	Score 35.6;	DB 20;	Length 2084;
	Best Local Similarity	64.6%;	Pred. No. 1.le+02;		
	Matches	53;	Conservative 0;	Mismatches 29;	Indels 0; Gaps 0;
Qy	19	CAAGAAAAAATCTCGAAAATTATTAAATATACATTTGATTTTTTATACAGTATT	78		
Dd	587	CAAGAAAAAGGACATGGTGAGTTATCAAAGGTGAACAGATTCATTTTGTACAGAACC	528		
Qy	79	ATAATGGAAGTAATCATGAGCCA	100		
Dd	527	ACAAGAAAGAAAGATGAGCCA	506		

## RESULT 8

```

US-10-433-793-138/C
; Sequence 138, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 138
; LENGTH: 5237
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-138

```

	Query Match	28.3%	Score 35.6	DB 19	Length 5237
	Best Local Similarity	67.6%	Pred. No. 1.4e+02		
	Matches 50	Conservative 0	Mismatches 24	Indels 0	Gaps 0
Qy	21	AGAAAAAATCCTCGAAAAATTATAATATACATTTTCATTTTATTTTATACAGTATTAT	80		
Db	3883	AAATAAAAAATCCACAAAAATCTTATAATTAATATTTTAAATTAATTTTATATAATAAC	3824		
Qy	81	AATGAGNACTACAT	94		
Db	3823	AAAAAAGTACTACGT	3810		

## RESULT 9

US-09-997-722-253  
; Sequence 253, Application US/09997722  
; Publication No. US20040072154A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David  
; APPLICANT: Enghard, Eric

```
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 253
; LENGTH: 27587
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-997-722-253

Query Match      28.1%; Score 35.4; DB 11; Length 27587;
Best Local Similarity 63.5%; Pred. No. 2.3e+02;
Matches 54; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 33 TCGAAATTTAATAATACATTTGATTTTATTTTATACAGATTTATAATAGAGAACTAC 92
Db 7439 TCAGAACTATATATATATATATTTTCTTCTACTAGCTGGGAACGCG 7498

Qy 93 ATGAGGCATCGGTGAGGGGAAC 117
Db 7499 CCAAGGAATTCATGTGTAGGAATC 7523

RESULT 10
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match      27.9%; Score 35.2; DB 16; Length 3673778;
Best Local Similarity 60.4%; Pred. No. 7e+02;
Matches 58; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 3 TAGCTTAGGAGGGCGCAAGAAAATCCCTCGAAAATTTAATAATACATTTGATTT 62
Db 1030773 TAGTTTAGGAGGAGGATAGGAGGAGTTATAGAGATTGTAATTTTGTATTT 1030832

Qy 63 ATTTTATACAGTATATATAGAGAACTACATGAGG 98
Db 1030833 TTTTTTTGTAGTATAATTTAATTTTATGAGG 1030868

RESULT 11
US-10-027-632-221497/c
; Sequence 221497, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221497
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-221497

Query Match      27.8%; Score 35; DB 17; Length 585;
Best Local Similarity 64.9%; Pred. No. 1.1e+02;
Matches 50; Conservative 1; Mismatches 26; Indels 0; Gaps 0;
```

```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221497
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-221497

Query Match      27.8%; Score 35; DB 13; Length 585;
Best Local Similarity 64.9%; Pred. No. 1.1e+02;
Matches 50; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

Qy 14 AGGGCGAAGAAAAAATCCCTCGAAAATTTAATAATACATTTGATTTTATTTATACA 73
Db 242 AGGGAAGGTAAAGAAAGATCAAAATACAGTAATTTGAATTTTGAATTTTATCT 183

Qy 74 GTATTATAATGAGAACT 90
Db 182 ATATCTCTGACAGCT 166

RESULT 12
US-10-027-632-221497/c
; Sequence 221497, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221497
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-221497

Query Match      27.8%; Score 35; DB 17; Length 585;
Best Local Similarity 64.9%; Pred. No. 1.1e+02;
Matches 50; Conservative 1; Mismatches 26; Indels 0; Gaps 0;
```



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 22, 2005, 03:58:26 ; Search time 3134 Seconds  
(without alignments)  
1530.344 Million cell updates/sec

Title: US-10-764-581-17  
Perfect score: 126  
Sequence: 1 catagacttagggggggca.....tgagggggaacatgatgaac 126

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsei:\*  
9: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.8	30.8	266	4	BM274717 PFEST0aa7
C 2	38.8	30.8	462	5	BP735538 BP735538
C 3	38.8	30.8	479	5	BP720714 BP720714
C 4	38.8	30.8	497	4	BJ629764 BJ629764
C 5	38	30.2	1101	9	CNS0127M AL108124 Drosophil
C 6	38	30.2	1134	9	CL110320 ISB1-53F3
C 7	37.2	29.5	792	7	CF284280 AGENCOURT
C 8	37	29.4	312	5	BP721218 BP721218
C 9	37	29.4	542	4	BG143667 ut58g07.x
C 10	37	29.4	635	9	CE073145 tigr-ges-
C 11	36.8	29.2	309	1	AV207641 AV207641
C 12	36.8	29.2	319	1	AV101294 AV101294
C 13	36.8	29.2	451	5	BU070396 im09f10.x
C 14	36.8	29.2	457	4	BM310085 ih11f02.y
C 15	36.8	29.2	459	7	H49659 yo23a07.r1
C 16	36.8	29.2	474	4	BM310349 ih11f02.x
C 17	36.8	29.2	534	5	BQ787408 im09f10.y
C 18	36.8	29.2	659	1	AL652174 AL652174
C 19	36.8	29.2	689	8	BZ322030 ia60c12.g
C 20	36.8	29.2	708	1	AL646926 AL646926
C 21	36.8	29.2	746	5	BX749863 BX749863
C 22	36.8	29.2	1011	9	CNS01420 AL104814 Drosophil
C 23	36.4	28.9	228	4	BJ338343 BJ338343
C 24	36.4	28.9	612	8	BH645122 BORUN84TR

25	36.4	28.9	1084	9	AG362103 Mus muscu
26	36.2	28.7	321	7	N36904 YY38a05.s1
C 27	36.2	28.7	454	9	CR314707 Medicago
28	36.2	28.7	623	4	BG602416 EST501506
29	36.2	28.7	756	8	BZ008596 Oe144C05.
30	36.2	28.7	782	8	BH440605 BOHNN36TR
31	36.2	28.7	853	9	CG971096 MBEDF69TR
C 32	36	28.6	566	6	CD268594 taag6C04..
C 33	36	28.6	732	9	BX186045 Danio rer
34	35.8	28.4	153	4	BJ393744 BJ393744
35	35.8	28.4	211	4	BJ332409 BJ332409
36	35.8	28.4	211	4	BJ393712 BJ393712
37	35.8	28.4	213	4	BJ332567 BJ332567
38	35.8	28.4	216	4	BJ331801 BJ331801
39	35.8	28.4	217	4	BJ367392 BJ367392
40	35.8	28.4	218	4	BJ334753 BJ334753
41	35.8	28.4	226	4	BJ393899 BJ393899
C 42	35.8	28.4	620	9	CR313043 Medicago
C 43	35.8	28.4	734	6	CB983990 AGENCOURT
C 44	35.8	28.4	735	6	CD811430 AGENCOURT
C 45	35.8	28.4	745	7	CF520593 AGENCOURT

## ALIGNMENTS

RESULT 1  
BM274717 266 bp mRNA linear EST 20-DEC-2001  
PFEST0aa73h03.y1 Plasmodium falciparum 3D7 gametocyte cDNA library  
Plasmodium falciparum 3D7 cDNA 5' similar to TR:077380 O77380  
PFC0780W PROTEIN. ; mRNA sequence.  
BM274717  
BM274717.1 GI:17968020  
EST.  
Plasmodium falciparum 3D7  
Plasmodium falciparum 3D7  
ORGANISM  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
REFERENCE  
1 (bases 1 to 266)  
AUTHORS  
Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D.,  
Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Thesling, B.,  
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I.,  
Tsagaris, V., Richey, J., Belaygorod, L., Franklin, C., Carr, L., Grow, A.,  
Maguire, L., Wadkins, J., Kennedy, S., Levinso, D.,  
Waterston, R., Wilson, R. and Sibley, D.  
WashU Plasmodium EST Project  
Unpublished (2001)  
Contact: L. David Sibley  
WashU Plasmodium EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
Library was constructed by R. Haywood. DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: L. David Sibley  
(sibley@borcim.wustl.edu), Washington University  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 237.  
Location/Qualifiers  
1..266  
/organism="Plasmodium falciparum 3D7"  
/mol\_type="mRNA"  
/db\_xref="taxon:36329"  
/dev\_stage="DH10B (GeneHog, Invitrogen, Inc.)"  
/lab\_host="DH10B (GeneHog, Invitrogen, Inc.)"  
/clone\_lib="Plasmodium falciparum 3D7 gametocyte cDNA  
library"  
/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:  
XhoI; The library was constructed by R Haywood. cDNAs were  
synthesized from gametocyte poly(A)+ RNA by oligo d(T)

67 TTATACAGTATTATATGAGAACTACATGAGGCATACGGGTGAGGG 112

SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
AUTHORS	1 (bases 1 to 457) Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.
TITLE	Expressed genes in <i>X. laevis</i> embryo
JOURNAL	Unpublished (2001)



```

SOURCE      Xenopus laevis (African clawed frog)
ORGANISM     Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE    1 (bases 1 to 792)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Daniela S. Gerhard, Ph.D.
              Office of Cancer Genomics
              National Cancer Institute / NIH
              Bldg. 31 Rm10A07 Bethesda, MD 20892
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Dr. Igor Dawid
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LAM12152 row: 1 column: 07
              High quality sequence stop: 680.
FEATURES     Location/Qualifiers
              1..792
              /organism="Xenopus laevis"
              /mol_type="mRNA"
              /db_xref="taxon:8355"
              /clone="IMAGE:5507094"
              /dev_stage="adult"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NICHD XGC Spl"
              /note="Organ: spleen; Vector: pCMV-SPORT6; Site 1: Not I;
              Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 2.4 kb. Constructed by Life
              Technologies."
ORIGIN
* Query Match      29.5%; Score 37.2; DB 7; Length 792;
Best Local Similarity 59.4%; Pred. No. 1.3e+02;
Matches 63; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 7 CTTAGGGAGGGCAAGAAAAAATCTCGAAAAATTATTAAATATACATTGATTTTATT 66
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 CTTAGTTGGTAAAGTCAAGCAACTGTTTATTATTACAGAGAAAAGAAATTA 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 67 TTATACAGTATTATATGAGAACTACATGAGGCATACGGGTGAGG 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 TTAATAATTTGTATTATATGATTATTAATGGGTCTATGGGATGG 353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
LOCUS      BP721218/c
DEFINITION BP721218 Osada Taira anterior neuroectoderm (ANE) PCS105 cDNA
            library Xenopus laevis cDNA clone XL448n2lex 3', mRNA sequence.
ACCESSION  BP721218
VERSION     BP721218.1
KEYWORDS    BP721218.1 GI:46069811
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE    1 (bases 1 to 312)
AUTHORS      Osada,S., Kitayama,A., Ueno,N. and Taira,M.
TITLE        Expression analysis of genes which are expressed in the anterior
            neuroectoderm of Xenopus embryos
JOURNAL      Unpublished (2004)
COMMENT      Contact: Masanori Taira
            Department of Biological Sciences
            Graduate School of Science, University of Tokyo; CREST, Japan
            Science and Technology Corporation, Japan

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: 81-03-5841-4434
Fax: 81-03-5841-4434
Email: m.taira@biol.s.u-tokyo.ac.jp.
URL: http://www.shigen.nig.ac.jp/nbrp/xenopus/est/.
Location/Qualifiers
1..312
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL448n2lex"
/tissue_type="anterior neuroectoderm"
/dev_stage="late gastrula (stage 12.5)"
/clone_lib="Osada Taira anterior neuroectoderm (ANE)
PCS105 cDNA library"
FEATURES     source
              1..312
              /organism="Xenopus laevis"
              /mol_type="mRNA"
              /db_xref="taxon:8355"
              /clone="XL448n2lex"
              /tissue_type="anterior neuroectoderm"
              /dev_stage="late gastrula (stage 12.5)"
              /clone_lib="Osada Taira anterior neuroectoderm (ANE)
              PCS105 cDNA library"
ORIGIN
Query Match      29.4%; Score 37; DB 5; Length 312;
Best Local Similarity 58.1%; Pred. No. 1.6e+02;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 8 TTAGGGAGGGCAAGAAAAAATCTCGAAAAATTTAAATATACATTGATTTTATT 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 TTAGTGNGATCAAGTNCAGGTACTGTTTAATTATTCAGAGAAAATGAATCAT 203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 68 TATACAGTATTATATGAGAACTACATGAGGCATACGGGTGAGG 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 TAAATAATTGGATTATTGATCATATAATGGAGTNTATGGAGACGG 158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
LOCUS      BG143667
DEFINITION BG143667 ut58g07.xl Soares mouse NMGB.bcell Mus musculus cDNA clone
            IMAGE:3332172 3', mRNA sequence.
ACCESSION  BG143667
VERSION     BG143667.1
KEYWORDS    BG143667.1 GI:12647066
SOURCE      Mus musculus (house mouse)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 542)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Other ESTs: ut58g07.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1076336
            High quality sequence stop: 470.
FEATURES     Location/Qualifiers
              1..542
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /clone="IMAGE:3332172"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="Soares mouse NMGB.bcell"
              /note="Organ: germinal B-cell; Vector: p7T3D-Pac
              (Pharmacia) with a modified polylinker; Site 1: Not I;
              Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
              oligo(dT) primer [5.
              TGTTCACCAATCTGAAGTGGAGCGCCGCTGGTGTGTGTGTGTGTGTGTGTGT
              T 3']; double-stranded cDNA was ligated to Eco RI
              adaptors (Pharmacia), digested with Not I and cloned into
              the Not I and Eco RI sites of the modified p7T3 vector.
              Library is normalized; constructed by Bento Soares and
              M.Fatima Bonaldo."
ORIGIN

```



Query Match 29.4%; Score 37; DB 4; Length 542;  
Best Local Similarity 67.5%; Pred. No. 1.5e+02;  
Matches 52; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

[illegible]

RESULT 10	CE073145/c	635 bp	DNA	linear	GSS 24-SEP-2003
LOCUS	tigr-gss-dog-17000323590267				
DEFINITION	Dog Library Canis familiaris genomic, genomic survey sequence.				

ACCESSION VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE AUTHORS
CE073145				
CE073145.1	GI:35136760			
GSS.				
		Canis familiaris (dog)		
		Canis familiaris		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
		1 (bases 1 to 635)		
		Kirkness, E.F., Batna, V., Halpern, A.L., Levy, S., Remington, K.,		
		Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and		
		Venter, J.C.		

**TITLE** The dog genome: survey sequencing and comparative analysis  
**JOURNAL** Science 301 (5841), 1898-1903 (2003)  
**MEDLINE** 22875432  
**PUBMED** 14512627  
**COMMENT** Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirknes@tigr.org  
 Class: shotgun.

```

FEATURES
source
name: 3
Location/Qualifiers
1. .635
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

```

ORIGIN		Percentage of Gaps	
Query Match	29.4%	Score 37;	DB 9; Length 635;
Best Local Similarity	64.7%	Pred. No. 1.5e+02;	
Matches	55; Conservative	0; Mismatches 30;	Indels 0; Gaps 0;
QY	21	AGAAAAAAAAATCCCTCGAAAAATATTAAATATACATTTGATTTTATTTTATACACAGTATTAT 80	
DB	207	ACAAAAAATGCGATGATGAATAAAATACAAATTTTATTTTATTTTATTTTATTAAGATTTT 148	

81 AATGAGAACTACATGAGGCATACGG 105  
147 ATTATTATTATTCATGAGACACAG 123

[illegible]

**KEYWORDS**  
**SOURCE**  
**ORGANISM**

REFERENCE  
AUTHORS

TITLE
JOURNAL
COMMENT

JOURNAL  
COMMENT

**FEATURES**  
**SOURCE**

## ORIGIN

Query Mat  
Best Loca  
Matches

.....

EST.  
Mus musculus (house mouse)

Mus musculus  
Eukaryota; Metazoa; Chorda  
Mammalia; Eutheria; Rodent  
1 (bases 1 to 309)

Konno, H., Aizawa, K., Akahira, F., Fukunishi, Y., H. Ito, Ishii, Y., Ishikawa, T., H. Kato, C., Kawai, J., Kikuchi, Matsuyama, T., Miki, R., Mizushima, Y., Ohsaka, Y., Owa, C., Ozawa, Y., Saico, H., Shibata, Y., Shigemoto, Y., Suzuki, H., Suzuki, H., Takata, T., Tsunoda, Y., Watahiki, A., W. Yokota, T., Yoshiki, A., Yoshida, K., and Y. Yoshida, RIKEN Mouse ESTs (Konno, H. Unpublished (1999))

Contact: Yoshihide Hayashi  
Laboratory for Genome Expl  
Sciences Center(GSC), Yoko  
The Institute of Physical  
1-7-22 Suehiro-cho, Tsurum  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp  
Sasaki, N., Izawa, M., Watahara, K., Matsui, T.,  
Matsuura, S., Carninci, P., Hayashizaki, Y.

Transcriptional sequencing  
polymerase. Proc. Natl. Acad.  
Itoh, M., Kitsunai, T. Aki

Tomaru, Y., Carninci, P., Sh  
Okazaki, Y. and Hayashizaki  
Automated filtration-base  
system. Genome Res. 9 (5),  
Carninci, P. and Hayashizaki  
High-efficiency full-leng  
19-44 (1999)

Location/Qualification  
1. .309

```
/organism="Mus mus  
/mol_type="mRNA"  
/strain="C57BL/6J  
/db_xref="taxon:1  
/clone="1700096D0  
/sex="male"  
/cissue_type="tes  
/dev_stage="adult  
/lab_host="SOLR"  
/clone_lib="PIVEN
```

```
/CONE_ID=-RIVEN
/note="Site 1: Xho
```

prepared and sequenced as part of the Genomic Sciences RIKEN Division project contributed to primed with a primer GAGAGAGAGAGATCCG prepared by using transcriptase and cap-trapper. Second primer adapter of GAGAGAGAGAGCGGGCAGG was cloned into pUC19.

2 : CENY : 1

### h

100

```

Qy 39 ATTATTAAATATACATTTGATTTTATTTTATATACAGTATTATTAATGAGAACTACATGAGG 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 ATTTAGTATATAAATTTGATTTTGTGTTTATACAGATATAAAGAATTCCTCATTAAT 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 99 CATACGGGTGAGGGGAACA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 CTTTCATGGGAAGGGAATA 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AV101294/c
LOCUS
DEFINITION AV101294 Mus musculus C57BL/6J ES cell Mus musculus cDNA clone
2410073H09, mRNA sequence.
ACCESSION AV101294
VERSION AV101294
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Naitsuma,H., Oda,H.,
Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
Sugahara,Y., Suzuki,H., Suzuki,H., Tateno,M., Tonaru,Y.,
Tomimaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
RIKEN Mouse ESTs
Unpublished (1999)
JOURNAL
COMMENT Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermotabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
    source
    Location/Qualifiers
        1..319
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="2410073H09"
            /cell_type="ES cell"
            /clone_lib="Mus musculus C57BL/6J ES cell"

ORIGIN
Query Match 29.2%; Score 36.8; DB 1; Length 319;
Best Local Similarity 66.2%; Pred. No. 1.7e+02;
Matches 53; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 39 ATTATTAAATATACATTTGATTTTATTTTATATACAGTATTATTAATGAGAACTACATGAGG 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 ATTTAGTATATAAATTTGATTTTGTGTTTATACAGATATAAAGAATTCCTCATTAAT 215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 99 CATACGGGTGAGGGGAACA 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 CTTTCATGGGAAGGGAATA 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
BU070396/c
LOCUS

```

```

DEFINITION in09f10.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6034699
3', mRNA sequence.
ACCESSION BU070396
VERSION BU070396.1 GI:22511585
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,D., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisshvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
JOURNAL
COMMENT Other ESTs: in09f10.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
University Genome Sequencing Center For information on obtaining a
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco.
FEATURES
    source
    Location/Qualifiers
        1..451
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6034699"
            /tissue_type="insulinoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="Human insulinoma"
            /note="Organ: pancreas; Vector: pBluescript SK-; Site:1;
            XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
            (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue following the Washington
            University protocol
            (http://genome.wustl.edu/est/lambda_protocol.shtml).
            Please contact Hiroshi Inoue, MD/PhD for further
            information on this library (Metabolism Division, Permutt
            Laboratory, Washington University School of Medicine, Box
            8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
            is a Washington University Pancreas EST project library."

ORIGIN
Query Match 29.2%; Score 36.8; DB 5; Length 451;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 5 GACTTAGGGGCGCAGAAAAAATCCTCGAAAAATTATTAAATATACATTTGATTTTAT 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 GAATTCCTCAAGGACAGAAAAAATGAGACTGTTGAATGCAGATTGAAGTAAT 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 65 TTTTATACAGTATTATATGAGAACTAC 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 TTTTAAATATATTATTTGGTTCGC 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
BM310085
LOCUS
DEFINITION ih1lf02.y1 Human insulinoma Homo sapiens cDNA 5', mRNA sequence.
ACCESSION BM310085
VERSION BM310085.1 GI:18042909

```

**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
**AUTHORS** 1 (bases 1 to 457)  
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blaisdell, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.  
**TITLE** Endocrine Pancreas Consortium  
**JOURNAL** Unpublished (2000)  
**COMMENT** Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 377.  
 Location/Qualifiers  
 1..457  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site: 1: XhoI; Site 2: EcoRI; Constructed with lambda ZAPRI system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda\_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."  
**FEATURES**  
 source  
 1..457  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site: 1: XhoI; Site 2: EcoRI; Constructed with lambda ZAPRI system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda\_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."  
**ORIGIN**  
 Query Match 29.2%; Score 36.8; DB 4; Length 457;  
 Best Local Similarity 63.6%; Pred. NO. 1.7e+02;  
 Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
 QY 5 GACTTAGGGGCGGCGAGAAAAATCTCGAAATATTAAATATACATTTGATTTAT 64  
 |||||  
 Db 297 GAATTCCTCAAGGGGACAGAAAAAATGGAGACTGTGGAATGCAGATTGAAGTAAT 356  
 |||||  
 QY 65 TTTTATACAGTATTATATGAGAACTAC 92  
 |||||  
 Db 357 TTTTAAAAATATTATTTGGGTTCTGC 384  
 |||||  
**RESULT 15**  
**LOCUS** H49659  
**DEFINITION** yo23a07.r1 Soares adult brain N2b5HB55Y Homo sapiens CDNA clone  
**IMAGE:** I78740 5', mRNA sequence.  
**ACCESSION** H49659  
**VERSION** H49659.1 GI:989500  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

**REFERENCE** Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
**AUTHORS** 1 (bases 1 to 459)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.  
**TITLE** The WashU-Merck EST Project  
**JOURNAL** Unpublished (1995)  
**COMMENT** Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 Insert Size: 442  
 High quality sequence stops: 327  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 442 Std Error: 0.00  
 Seq primer: M13RP1  
 High quality sequence stop: 327.  
 Location/Qualifiers  
 1..459  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:3840936"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:178740"  
 /sex="Male"  
 /dev\_stage="55-year old"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares adult brain N2b5HB55Y"  
 /note="Organ: brain; Vector: p77T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCCGCTTTTCTTTTCTTTTCTTTT 3'], digested with Not I and Eco RI sites of a modified p77T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M. Fatima Bonaldo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, subcortical white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and medulla."  
**ORIGIN**  
 Query Match 29.2%; Score 36.8; DB 7; Length 459;  
 Best Local Similarity 63.6%; Pred. NO. 1.7e+02;  
 Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
 QY 5 GACTTAGGGGCGGCGAGAAAAATCTCGAAATATTAAATATACATTTGATTTAT 64  
 |||||  
 Db 116 GAATTCCTCAAGGGGACAGAAAAAATGGAGACTGTGGAATGCAGATTGAAGTAAT 175  
 |||||  
 QY 65 TTTTATACAGTATTATATGAGAACTAC 92  
 |||||  
 Db 176 TTTTAAAAATATTATTTGGGTTCTGC 203  
 |||||  
 Search completed: July 22, 2005, 08:38:52  
 Job time : 3142 secs

***This Page Blank (usp19)***